



U.S. Department
of Transportation
**National Highway
Traffic Safety
Administration**

**DOT HS 807 862
Review of the Literature**

March 1992

The Deterrent Capability of Sobriety Checkpoints: Summary of the American Literature

This publication is distributed by the U.S. Department of Transportation, National Highway Traffic Safety Administration, in the interest of information exchange. The opinions, findings and conclusions expressed in this publication are those of the author(s) and not necessarily those of the Department of Transportation or the National Highway Traffic Safety Administration. The United States Government assumes no liability for its contents or use thereof. If trade or manufacturers' name or products are mentioned, it is because they are considered essential to the object of the publication and should not be construed as an endorsement. The United States Government does not endorse products or manufacturers.

1. Report No. DOT HS 807 862		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle The Deterrent Capability of Sobriety Checkpoints: Summary of the American Literature				5. Report Date March, 1992	
				6. Performing Organization Code	
7. Author(s) H. Laurence Ross				8. Performing Organization Report No.	
9. Performing Organization Name and Address Department of Sociology University of New Mexico Albuquerque, NM 87131				10. Work Unit No. (TRAI5)	
				11. Contract or Grant No. DTNH22-91-Y-05317	
12. Sponsoring Agency Name and Address National Highway Traffic Safety Administration Office of Alcohol and State Programs, NTS-20 400 Seventh St. SW Washington, DC 20590				13. Type of Report and Period Covered Review of the Literature 1980 - 1991	
				14. Sponsoring Agency Code NTS-20	
15. Supplementary Notes					
16. Abstract This report reviews and evaluates the scientific literature on sobriety checkpoints in the United States. Concerns about the constitutionality of checkpoint procedures initially limited the number of checkpoint programs in this country as well as constraining the procedures used. However, nine case studies were located, along with a multi-state econometric study of various drunk-driving laws including those authorizing checkpoints. Although due to methodological problems no one study is convincing, the accumulation strongly supports the proposition that sobriety checkpoints can deter impaired driving. A review of selected foreign experience also supports this proposition. The totality of the evidence suggests that the following factors are important in maximizing the deterrent impact of checkpoints: (1) the clarity of purpose with which they are conducted; (2) the frequency of the checkpoints; (3) media attention; and (4) positioning in a diversified program of law enforcement related to impaired driving.					
17. Key Words Sobriety Checkpoints; deterrence; drunk driving; law enforcement; effectiveness			18. Distribution Statement Document is available to the U. S. public through the National Technical Information Service, Springfield, VA 22161		
19. Security Classif. (of this report)		20. Security Classif. (of this page)		21. No. of Pages	22. Price

TABLE OF CONTENTS

Executive summary iii

Introduction 1

Methodological considerations 2

 Measurement issues 3

 Interpretation issues 3

Scientific evaluations of sobriety checkpoints 5

 The U.S. literature 5

 Relevant foreign studies 10

Conclusion 14

References 15

Table 1 19

Figure 1 21

EXECUTIVE SUMMARY

This report reviews the evaluation literature concerning police sobriety checkpoints in the United States. It reaches the conclusion that, although most of the studies contain important methodological weaknesses, the cumulation of evidence supports the hypothesis that checkpoints reduce impaired driving. This is derived from theoretical reasoning that the key element in deterrence is the perceived likelihood of punishment for impaired drivers, and that checkpoint programs, by multiplying the occasions of interaction between the driving public and law enforcement personnel and by liberating this interaction from a link to manifest driving errors, foster this perception.

The literature is relatively sparse and most of the interventions are limited and cautious, reflecting the existence of doubts concerning checkpoints' constitutionality prior to the 1990 Sitz case, in which the Supreme Court determined that they did not violate the U.S. Constitution. Moreover, the literature as a whole suffers from common methodological problems in the areas of measurement and inference.

A serious weakness is the use of inadequate measures of impaired driving, most notably alcohol-related accidents as defined by the police. This index has a large subjective component, and the police decision process is potentially subject to political influence. Even more problematic are those studies that use self-

reported drunk driving from telephone surveys as their sole or chief measure. Furthermore, the interventions themselves tend to be badly described, and they are often complex, involving numerous components other than checkpoints which may independently or in interaction with checkpoints be responsible for changes that are attributed to the checkpoints alone.

Another consideration is problems of study design. Although checkpoint campaigns, like other enforcement interventions, are generally expected to produce immediate effects and therefore can be analyzed by interrupted time-series methods, several of the studies employ inherently defective before-and-after comparisons with weak or nonexistent controls. Alternative explanations in terms of independent events, general trends, random fluctuations, and the tendency of unusual conditions to return to normal, exist as threats to the validity of conclusions in these studies.

However, the weaknesses of individual studies appear to be adequately overcome by the accumulation of positive findings for visible and well-publicized checkpoints. The accomplishments have been demonstrated only in the short run, mainly because most of the programs were either new when evaluated or were deliberately limited in time. The only evidence of relatively long-term deterrence from something like checkpoints comes from New South Wales, Australia, and that intervention is both quantitatively and qualitatively different from typical U.S. checkpoint programs. The Australian experience can be regarded as suggestive but not

demonstrative of what can be achieved by the kinds of programs experienced here.

It can be concluded that sobriety checkpoints are capable of reducing the prevalence of drunk driving. However, more must be learned concerning the conditions under which this capability is realized. Among the plausible considerations are:

1. The nature of the checkpoint program, especially the degree to which the purpose of deterring drunk driving is made evident to the public.
2. The frequency of checkpoints and of driver contacts.
3. The amount and nature of publicity received.
4. The policy context in which the checkpoints are embedded, including other drunk driving law enforcement.

THE DETERRENT CAPABILITY OF SOBRIETY CHECKPOINTS:
SUMMARY OF THE AMERICAN LITERATURE

H. Laurence Ross

Introduction

Sobriety checkpoints are a theoretically promising tool for deterring people from driving while intoxicated. This report evaluates the scientific evidence of deterrent effectiveness of sobriety checkpoints in the United States as of 1992. For each reported study it summarizes the intervention and the analysis used by the evaluators to determine the consequences. It assesses the persuasiveness of each discrete study and speculates concerning the implications of the totality of available knowledge for deterrence theory and social policy.

The theoretical advantages of sobriety checkpoints over other law enforcement techniques relate to the fact that the objective risk of apprehension for driving while intoxicated is extremely low -- about 1 arrest to 1000 occasions. This vitiates the effect of the deterrent threat posed by drunk-driving law. Checkpoints aim both to increase the objective risk and, especially, to affect the subjective risk of apprehension for a drinker considering driving. They are expected to do so because they create numerous contacts between the police and the driving public and are also evident to large numbers of passers-by, for example drivers in the lane opposite to that blockaded by the checkpoint. In addition, partly because of their novelty in the United States, they are attractive to the media, thus spreading the message of law enforcement to drivers who may not have direct contact with the checkpoints. Furthermore, they can subvert the technique of neutralization whereby a drinker can reason that his superior driving performance while impaired will protect him from police attention:

Drivers may believe that they stand little chance of being detected if they drive after drinking too much. They may believe that the police will not notice them or that they can drive carefully enough not to attract suspicion if they are noticed. But roadblocks, or safety checkpoints, counter this belief because the potential for a drinking driver to be detected is increased (Dickman, 1987:2).

The issue for evaluation concerns how well these theoretical advantages have been realized in practice. Despite the deterrent promises made by proponents of sobriety checkpoints, many U.S.

police agencies appear diffident concerning their use. Although the reasons for this have not been thoroughly explored, among them are very likely the competition with other police functions for resources, especially on weekend nights, fear of damage to public relations, and the belief that the proper criterion for police performance is arrests of actual criminals rather than deterrence of potential criminals. The research record may be of use in suggesting arguments to influence a skeptical police audience of the validity of deterrence theory for their practical concerns.

This literature also bears implications for the acceptability of sobriety checkpoints to the public and to the courts. Like all police actions, checkpoints consume personnel and equipment resources. This occurs in competition with many other police functions, such as responding to requests for assistance from citizens. Furthermore, checkpoints have been criticized as intrusive, constituting seizures for the purpose of investigating criminal activity in the absence of prior suspicion that a crime has been committed and that a particular individual is involved. These costs in terms of liberties and resources can be justified by benefits in terms of effective and efficient deterrence of impaired driving (Michigan Department of State Police v. Sitz, 1990). Until recently, the favorable balance has been a reasonable assumption, but one based more on conjecture than persuasive evidence. The costs of checkpoint programs could be most easily accepted on the basis of the results of competent evaluations testifying to their deterrent accomplishments.

Methodological considerations

Appraisal and interpretation of the research literature requires understanding of the nature of each intervention, that is, the type of activity, its intensiveness in terms of numbers of checkpoints and numbers of drivers stopped in relation to the target population, its extensiveness or duration, and the extent of publicity. Surprisingly, a large proportion of reports are uninformative on these matters, although existing checkpoint programs vary greatly. For example, the numbers of checkpoints conducted for the intervention vary between 2 in the Arizona report by Epperlein (1985) to more than 100 in Charlottesville (Voas et al., 1985) and, presumably, Bergen County, New Jersey (Levy et al., 1990) compared with many thousands annually in New South Wales, Australia (Homel, 1990). Perhaps the least adequate information concerns the publicity accompanying the enforcement effort, partly because overall publicity is not easily quantified. Most of the programs described appear to have been novel and to have received considerable news coverage, especially in smaller cities, but in some cases, where the results have been less than expected, publicity seems to have been lacking.

Measurement issues. There are a variety of measures of drunk driving available which have been used to determine the effectiveness of interventions. The most direct is blood-alcohol concentrations (BACs) in the driving population. Although this is usually considered an expensive measure to obtain, checkpoints themselves can often provide the data. Ideally, checkpoint data should be compared with pre-intervention baseline data from research roadblocks.

As supplements or substitutes for a direct measure of impaired driving, surrogate indexes such as nighttime fatalities, single-vehicle fatalities, or even total fatal and serious-injury crashes can be useful. Caution is needed in interpreting these, because they also reflect the influence of factors other than alcohol. Consequently, the directness of the index may change over time, leading to erroneous impressions of reduced impairment. However, from a pragmatic perspective reducing fatalities and serious injuries is in itself a legitimate goal, and any program that efficiently and effectively lowers fatality and injury rates, whatever the mechanism, is certainly defensible.

Another good surrogate measure is the proportion of crash fatalities associated with alcohol, which is obtained from data bases like the Fatal Accident Reporting System (FARS). When interpreting studies using this index, the possibility of differential reporting and of changes in the causal relationship between alcohol impairment and traffic fatalities over time must be kept in mind.

In contrast, some other surrogate measures in the evaluation literature are inherently less valid and exclusive dependence on them compromises the usefulness of some studies. As an extreme example, arrests are at best weakly related to the amount of impaired driving and are not acceptable as a legitimate index. None of the studies considered here use arrests as an index of drunk driving, but several use alcohol-related crashes where the agency reporting the alcohol is the same as the one implementing the intervention. Because decisions concerning the role of alcohol in accidents are to an important degree subjective and possibly subject to political influence, these data and the conclusions derived from them should be regarded with skepticism. Self reports of driving after drinking should also be considered potentially unreliable, especially as the current social concern about drunk driving may change the willingness of people to report it independently of any change in their behavior.

Interpretation issues. In addition to these measurement problems, there are numerous problems of causal interpretation characteristic of the circumstances surrounding sobriety checkpoint demonstration programs. In some cases they considerably weaken the confidence that can be placed in the

claimed results. The most common can be listed as follows:

1. Other relevant events, specifically including additional drunk-driving countermeasures occurring simultaneously with the checkpoints, may explain observed changes in the amount of drunk driving. The Clearwater-Largo study (Lacey et al., 1986) is greatly compromised by this factor as evidence concerning the effectiveness specifically of checkpoints.

2. Trends, both secular and cyclical, may be responsible for the changes. The existence of cyclical trends means that interventions of relatively short duration may have their results skewed by, for example, a seasonal effect; casualties are generally lower in November than in August, so interventions in September and October will appear more successful than they really are. The Arizona intervention (Epperlein, 1985) was conducted just before Christmas and could have faced this problem. Longer-range trends can also give a false impression of success; this affects the interpretation of studies that compare post-intervention years or averages with previous ones. Some otherwise impressive sobriety checkpoint studies, including the Charlottesville report (Voas et al., 1985), share this problem.

3. Random variation may yield unwarranted impressions of effect. Especially in small jurisdictions this variation can be great enough to hide real intervention effects. The problem is greatest when the measure uses uncommon events like fatalities as a base. Tests of statistical significance are designed to minimize the chances of perceiving a causal relationship when none exists. Significance testing is general, but not universal, in this literature. The probability of the opposite error, concluding no effect when there is one, can be calculated by power analyses, which are seldom mentioned in these publications.

4. Regression of extreme measures may be mistaken for an intervention effect. If drunk driving countermeasures are conducted, and the data measured, at places with unusually high drunk-driving statistics, an impression of success can be favored by the general tendency of extremes to return to averages. Such selection is common in specifying the sites for sobriety checkpoints, for example in Bergen County (Levy et al., 1990), and it could present an interpretation problem in the absence of comparable controls.

Table 1 presents a summary description of the recent evaluation literature concerning sobriety checkpoints in the United States, noting the methods and conclusions. The following section of the report describes these studies in somewhat more detail and suggests their strengths and weaknesses. A further section adds material from selected foreign studies, and the conclusion evaluates the evidence for checkpoint effectiveness taken as a whole. It finds that although most of the studies

reported in the literature to date contain methodological weaknesses, the cumulation of evidence in the literature as a whole provides a reasonable basis for concluding that appropriately mounted sobriety checkpoint campaigns are capable of deterring drunk driving.

Scientific evaluations of sobriety checkpoints

The U.S. literature. The available evaluations of American experience with checkpoints are presented here, as in Table 1, in chronological order. As the use of sobriety checkpoints is relatively new in the U.S., it is not surprising to find that reported evaluations go back only to 1984.

The first relevant study was the Williams and Lund (1984) telephone survey of opinions about checkpoint use in Maryland, Delaware, and Virginia, where residents of counties in which checkpoints were held were compared with residents of similar counties lacking checkpoints. Although survey data standing alone cannot be considered powerful evidence of behavior, as noted previously, the comparison of residents of counties on the Delmarva Peninsula is enlightening because the Delaware police engaged in a checkpoint program of considerable intensity, considering the relatively sparse population of that state's two southern counties. The comparison of Montgomery County, Maryland, using checkpoints, with Fairfax County, Virginia, which did not, is enlightening because the drunk-driving arrest rate was objectively much higher in Fairfax. The study found that checkpoint programs were associated with significantly increased public awareness of police activities, including the impression of a relatively higher chance of arrest for an impaired driver. Both Montgomery and Fairfax residents contrary to fact perceived the chance of arrest for a drunk driver to be higher in Montgomery. Similar but less pronounced results were found for residents of the Delmarva Peninsula. Disappointingly, this greater risk perception was not associated with reduced self-reported impaired driving (the study did not employ other types of impaired-driving measures). The situation is reminiscent of that in New Philadelphia, Ohio, where the well-publicized activities of the local judge resulted in greater perceived certainty as well as severity of punishment compared with a quasi-experimental control community, yet no evidence was found of less impaired driving (Ross and Voas, 1989). A possible explanation of this disconfirmation of expectations is that the perceived risk fell short of a necessary threshold. Although Delaware's program was intensive in the general American context, checkpoints held only weekly in a fairly large area may be too few to be effective as threats, albeit sufficiently numerous to gain attention. The Montgomery County program appears to have been even more diluted, given the larger population and roadway network.

The Arizona study (Epperlein, 1985) is remarkable for its elegant interrupted time-series evaluation method, but the stimulus -- two checkpoints on one weekend in each of three sites -- was merely token. That an impact (a statistically significant decline of 28 percent in the area's serious crash ratio) was found is surprising in light of the limited nature of the intervention, but that it disappeared almost immediately when the checkpoints were withdrawn is not.

The Charlottesville study (Voas et al., 1985) is in some ways the most informative of the American evaluations. The intervention was relatively intense, amounting to nearly two operations a week on the average, in a small city where word of mouth might be relied on to spread the news about police activity. The study used a wealth of approaches to determining effects and employed sophisticated controls, both a specific comparison community and the state of Virginia as a whole. As in Maryland and Delaware previously, surveys found that the checkpoints were noticed by the public. In addition, predicted reductions were found in six different comparisons using surrogate field measures of impaired driving. The report focused on the fact that the program was associated with a 13 percent reduction in the proportion of alcohol-related accidents. However, the Charlottesville study fails to provide the clear support for the deterrent capability of sobriety checkpoints that one might hope for, given its repeated citation as evidence of success. One problem lies in that the survey data did not find that the perceived risk of punishment had increased in Charlottesville over time, either in absolute terms or in comparison with the control community. Also bothersome is that the method, comparing conditions during the checkpoint program with an average level for the previous three years, fails to control plausible alternative explanations related to possible trends in the data. Moreover, the changes in the field measures, although all in the expected direction, were not statistically significant in three of six comparisons, and the apparent effect was greater using the suspect "alcohol-reported" measures than with the less controversial nighttime crash index. The Charlottesville results are encouraging but they are far from being definitive evidence of sobriety checkpoint effectiveness, even in a relatively intensive application.

The Clearwater-Largo study (Lacey et al., 1986), like Charlottesville, documents changes in public perceptions as well as in field data that index impaired driving. Moreover, unlike Charlottesville, this study found increased perceived risk of apprehension and punishment, and the analytical method applied to the field data was interrupted time series, methodologically superior to Charlottesville's before-after comparison. A 20-percent decline was found in alcohol-related accidents, unfortunately not the best criterion variable as explained above. The study was competently executed, with appropriate controls.

Significant reductions were found in both alcohol-related crashes and nighttime crashes. The chief interpretive problem in dealing with Clearwater-Largo as a demonstration of checkpoint effectiveness is that the intervention was complex, including enhanced traditional patrol at critical locations, training of specialized police officers, and improved DWI processing procedures, among other components. Moreover, the checkpoint component appears to have been minimal, a total of 12 taking place in 18 months. Although these were described as "media events" and were cited by survey respondents in describing what was new about drunk-driving enforcement, the conservative conclusion would be that the complex of programs constituting the intervention had deterrent results but that the contribution of each component -- including checkpoints -- must be regarded as indeterminate. If the sobriety checkpoints in Clearwater-Largo "worked," it may have been because they were accompanied and reinforced by other aspects of the intervention. Indeed, the investigator themselves attributed the reduction in alcohol-related crashes to "the combined DWI enforcement and public information program in Clearwater/Largo" (Lacey et al., 1990: 32).

The interpretive problems of Clearwater-Largo are even more bothersome in the Indianapolis study (Lacey et al., 1988), which may be one reason why it is only infrequently cited in materials on the deterrent effectiveness of sobriety checkpoints. While the Florida communities launched a mere 12 checkpoints, Indianapolis had just half that number. These were noted by respondents in surveys, but there was no evidence of significant increases in the perception of risk of punishment for impaired drivers and the interrupted time-series analyses failed to find significant changes in the indexes of impaired driving. This is of course another reason for the study's being overlooked. However, given the limited nature of the intervention, expectations of major deterrent accomplishments were probably unwarranted from the beginning.

Hingson and Howland's (1989) telephone survey of Massachusetts residents is of limited value, in part because the report does not describe the intervention in any detail. Massachusetts police use checkpoints and those in New Hampshire, due to an adverse state supreme court decision, cannot. However, the nature and numbers of the Massachusetts checkpoints are not reported. Furthermore, the results are based on respondents' opinions only, uncorroborated by field data like nighttime traffic fatalities. As predicted by theory, and in contrast to the findings of the Williams and Lund study reported above, respondents in Massachusetts report less driving after drinking than do those in the comparison jurisdiction, but the plausibility of other possible causes beyond the checkpoints is obvious. For example, New Hampshire police may also have a very low overall drunk-driving arrest rate. Perhaps New Hampshire

residents, induced by the cheap liquor sold in the state monopoly stores, drink more in general than do people who live in Massachusetts.

The Bergen County study (Levy et al., 1988; 1990) evaluates one of the more intensive and sustained sobriety checkpoint programs in the U.S. literature. Although the reports do not indicate the number of checkpoint operations, they refer to around 35,000 driver checks annually over the course of 2 1/2 years. Other, overlapping, programs included an educational program and a community DWI task force, and these may well have obtained the attention of local media to publicize the checkpoints. However, a handicap for the Bergen County program is that the principal media in the county serve the entire New York metropolitan area, a context in which something like suburban sobriety checkpoints may not have a strong claim for attention. The analysis of the program's impact relies on econometric modeling of total reported single-vehicle nighttime accidents along with a variety of quantitative independent and control variables. This type of analysis is reputed to lead to robust findings in terms of the direction of relationships, but not necessarily in terms of their degree. The authors found that a statistically significant decline of 10 to 15 percent was experienced in the criterion variable, with a peak effect in the second year and a subsequent decline. Although the analysis is sophisticated and the index is a reasonable one, the work is not as conclusive as one might hope. For one thing, the finding of deterrent effectiveness for the checkpoint program was not replicated with single-vehicle nighttime fatalities, generally considered a better index of alcohol-impaired driving. For another, the concurrent experience of two other programs complicates the interpretation of the results. Indeed, the authors make the modest claim "that single vehicle night-time fatalities have declined as DWI programs have been introduced, but do not permit us to impute a defined proportion of the declines to specific programs" (1990, pp. 253-254). Furthermore, as in most analyses of this kind, it is not evident that all relevant contextual and competing causal variables were properly accounted for.

A recent and relatively persuasive analysis of the deterrent consequences of sobriety checkpoints is provided in a study by the Insurance Institute for Highway Safety of an enforcement program in Binghamton, NY (Wells, et al., 1991). The intervention was not on the scale of Charlottesville or Bergen County, but it was substantial, involving 72 checkpoints in two years, in which nearly 10,000 drivers were tested in this city of around 50,000 people. Both impaired driving and seat belt use were the declared subjects of the highly publicized checkpoint operations. The analysis employed survey data, finding that, at least initially, awareness of DWI enforcement rose considerably in Binghamton but not in a comparison city, and that the awareness centered on

checkpoints. Perceptions of the risk of punishment rose at the program's inception, though one can infer that they subsequently declined from their initial peak. The checkpoints yielded evidence that there occurred a nearly 40 percent decline in drivers who had been drinking, and this decline was sustained through the second year of the program. Moreover, there were statistically significant declines in serious crashes and in late-night crashes during the checkpoint months (none were held in summer or mid-winter). The main interpretive problem comes from the finding that there was little change in the numbers of drivers with importantly elevated blood-alcohol concentrations (over 0.10% BAC), who constitute the core of the drunk-driving fatality problem. The possible resistance of the most committed deviants to deterrent threats, suggested in the general criminal literature for some time (e.g., Zimring and Hawkins, 1973), may explain this finding.

In contrast to most of the examples considered so far, a study by the office of the New Mexico State Epidemiologist of the results of four checkpoint campaigns on that state's highways appears to be yielding negative results. (Information was obtained in private communications from Stewart Castle.) The most intensive and extended campaign site included 101 checkpoints over two years, with 30,000 tests, a considerably larger effort than Binghamton, concentrated at a single location instead of being spread across a road network. Comparing fatal and injury crashes at all sites with pre-program levels, an 11 percent decline was found in the first program year but there was an 11 percent increase in the second. Because the intervention in two of the sites was modest, terminating after the first year, and a third site faced a seemingly intractable population on the Navajo reservation and was also cut short, it is reasonable to limit concern to the principal site. However, these results were also disappointing -- a 14 percent decrease in serious crashes in the first year but a 21 percent increase in the second. These crude before-and-after comparisons may not be definitive, but it is unlikely that data will be available to permit a more sophisticated analysis and the preliminary results do not encourage faith that more positive findings would emerge from one. The apparent failure of sobriety checkpoints to deter drunk driving in the New Mexico study may well be a consequence of a lack of publicity. The checkpoints were focused on road segments, not communities with local media. Therefore, they may have been both relatively uninteresting as local news and minimally reported by word of mouth. The evaluators relate that there was very little attention in the state media after the checkpoints were launched.

The issue of sobriety checkpoint effectiveness has also been addressed in an econometric analysis of multi-state data by Evans et al. (1991). The data base of this study included information on a variety of statutory approaches to reducing drunk driving in

the 12-year period 1975-1986: administrative per se laws, prohibitions against plea bargaining, laws mandating jail or community service on the first offense, illegal per se laws, open container prohibitions, laws authorizing pre-arrest breath tests, and those authorizing sobriety checkpoints. Previous analyses by others using similar data had found significant impacts for illegal per se and administrative per se laws and, in one instance, first-offender jail or community service (Zador et al., 1988; see also Joksch, 1988, and Klein, 1989). However, in this most comprehensive analysis, controlling for unemployment, vehicle miles traveled, beer taxes, and mandatory belt-use legislation, Evans et al. found none of the previously noted variables to be related to measures of impaired driving. Rather, among these deterrence-based laws, only those authorizing sobriety checkpoints appeared to have a significant deterrent effect on drunk driving. The authors adhere to the view that perceived certainty of punishment is the most effective of the deterrence variables, and they interpret the results as evidence that authorizing checkpoints leads to maximizing the perceived certainty of punishment for drunk drivers.

Relevant foreign studies. Because many countries lack equivalents to the constitutional constraints on American criminal procedure, jurisdictions outside the U.S. initiated sobriety checkpoint programs considerably earlier, yielding a literature that furthers understanding of checkpoints' potential capabilities and limitations. These earlier studies are reviewed in my book, Deterring the Drinking Driver (Ross, 1982). Care must be exercised in generalizing from other countries to the American scene, for two reasons. First, although the Supreme Court has approved the use of checkpoints, they are conducted under more constraints in the U.S. than in many other places. American police may stop all passing drivers or a systematic sample of them in checkpoints, but not merely some few selected at the officer's discretion. Moreover, the police must develop suspicion concerning a specific individual driver before requiring participation in a breath test for alcohol. The use of passive alcohol sensors in screening stopped drivers may eventually reduce the proportion of drinking drivers who fail to be apprehended during the brief checkpoint interview, but demanding active participation in breath tests of all passers-by, as is done in Scandinavia and Australia, would be of doubtful constitutionality in the United States. Second, cultural differences among technologically developed countries are strong in such matters as the place of alcohol in the society and the available transportation opportunities. A technique that works well in Sydney or Helsinki might prove disappointing in Washington or Albuquerque because of such differences. With these cautions in mind, foreign literature can augment the U.S. findings.

Foreign literature offers several examples of enforcement

campaigns embodying sobriety checkpoints which importantly reduced the amount of impaired driving, even in countries with high alcohol consumption and where concentrated binge drinking occurs. However, the accomplishments have usually diminished shortly following termination of the campaigns. Among the most interesting examples are the study by Cameron, et al. (1980) in Victoria, Australia, and that by Ross et al. (1982) in France. The Victorian experience is of note partly because the activities there provided a model for the subsequent more intensive and prolonged intervention in neighboring New South Wales. France is of interest because the constraints on the planning and execution of checkpoints at the time -- announcement in advance, providing a show of force on the scene, etc. -- were similar to those held generally applicable in the U.S. following the 1990 Sitz case (NHTSA, 1990).

The Victorian "blitzes" in 1977 and 1978 were among the earliest enforcement campaigns employing sobriety checkpoints. The evaluation methods applied would not be considered adequate today, but they provide arguably persuasive evidence that the blitzes raised public perception of police activity and of the risk of an impaired driver's being caught, and that both nighttime serious crashes and the BACs of drivers in single-vehicle crashes diminished. No measures were done following termination of the campaigns to determine the duration of these effects.

The French study used the results of existing surveys and applied interrupted time-series methods to official crash statistics. It found that, in part due to serious principled opposition, the law's provisions were recognized by 97% of the adult population, a record for knowledge of legislation in France and, perhaps, anywhere. There was a 12 percent decline in crash-related injuries at the law's inception, but the gain was 95 percent dissipated within 9 months. There was also a 14 percent decline in crash-related fatalities, 95 percent of which was gone in 13 months. The findings of initial effectiveness were bolstered with comparisons between weekday and weekend night crashes and between crashes in parts of the country differing in alcohol consumption.

Foreign literature also demonstrates that checkpoint enforcement requires supplementation with publicity in order to be transformed into perceived punishment risk, which is the theoretical requisite for deterrence. The most important study is that of Mercer (1985). He had the mixed fortune to study a massive checkpoint-based enforcement campaign in British Columbia, Canada at a time that included a strike in the province's major news media. The intervention consisted of a month-long blitz in which 265,000 vehicle stops were made without any apparent impact on impaired driving. Surveys before and after the blitz showed that the public was largely unaware of it, even

though 17 percent had seen a checkpoint and 6 percent had been tested. Interrupted time-series analysis showed no change in alcohol-related injurious crashes. Moreover, a correlational analysis using five years' worth of provincial data led to the conclusion that, "when the effects of the other variables are controlled for, media coverage was the only statistically significant predictor of a reduction in either the number or percentage of alcohol-related casualty accidents" (p. 472). Numbers of checkpoints and numbers of arrests were unrelated to casualties when other variables were controlled. It is unfortunate that Mercer used only one, relatively weak, index of impaired driving, but the study can be viewed as at least moderately convincing in light of his general (1984) finding of effectiveness for checkpoint operations.

The most important indication of the deterrent potential of policies similar to sobriety checkpoints comes from the experience of "random breath testing" (RBT) in the Australian state of New South Wales. One must note that RBT in New South Wales is not the same as U.S.-style sobriety checkpoints. The enabling legislation empowered police to demand a breath test of a driver without the need for individualized suspicion of impaired driving. In New South Wales checkpoints, all drivers are routinely breath-tested on a portable instrument. Moreover, the government initiated RBT with unprecedented vigor and publicity. The first year's figure of a million tests in a population of 3 million drivers dwarfs prevailing programs in the U.S. and most other countries (though not Finland, where the checkpoints have not been scientifically evaluated). Public funds paid for advertising including catchy jingles that saturated television at the time RBT was launched. The program has been described and evaluated in numerous publications, many from around 1985 undertaken in order to secure a permanent status for what was enacted as temporary policy (Arthurson, 1985; Carseldine, 1985; see also Homel, 1988).

Random breath testing in New South Wales was not only introduced "boots and all," in Homel's words, but it was extremely fortunate in its temporal coincidence with an economic recession. Modeling of the intervention and economic variables by Thomson and Mavrolefterou (cited and reprinted in NRMA, 1985) suggested that RBT was responsible for less than half the fatality savings in the first year, although somewhat more than half in the second. The public thus were led to believe in an effect of RBT that was on the order of twice its actual causal impact on fatalities. The practical import of this is that RBT quickly became viewed as an effective lifesaving intervention. With its effectiveness established, public support reached near-unanimous levels, and initial police reticence concerning activity that seemed to yield fewer arrests and more public relations problems than standard patrol was largely overcome. This facilitated further enforcement and, presumably, further

deterrence.

Perhaps the most important claim made for Australian RBT is that the impact, on the order of a 20 percent decline in total crash fatalities, is possibly permanent. Analyses of experience during the first three years gave a close fit to a model of a sustained reduction in the time series of fatal crashes. Figure 1, from Arthurson (1985), illustrates this. However, some interpretive cautions have been raised in more recent years (NRMA, 1991). First, although it is less than before RBT, the role of alcohol in fatal and serious crashes remains large. Of drivers and motorcyclists killed in New South Wales, 42 percent had illegal BACs (0.05% and higher) before 1983. This dropped to 36 percent in 1983, and 33 percent in 1984, with little progress thereafter. Second, recent surveys report the public impression, apparently inaccurate, that there is a reduced level of RBT activity. This may be due to the fact that some proportion of RBT since 1987 has been "mobile," that is, done not in checkpoints but rather by individual police units. However, in 1990, "stationary" RBT -- sobriety checkpoints -- administered nearly 1.5 million breath tests, compared with less than a quarter-million done in the mobile mode (RTA, 1991). There were 50 percent more tests by police in the stationary mode in 1990 than in 1983. Another reason offered for the prevailing impression is that the stationary units appear less obvious than in the past, due to the police decision to withdraw the large "booze buses" which figured prominently in the early days of RBT in favor of cars with simpler identification in the form of a sign obtained by lifting the lid of the trunk. Even the latter has become rare because recent New South Wales patrol cars have a different design and are not capable of displaying the signs. These considerations have led the NRMA (a group comparable to the American Automobile Association) to suggest a "relaunching" of RBT with renewed publicity.

In summary, RBT does appear to have made good on its promise of traffic safety, and most indications point to an extraordinary length as well as depth of effect. Compared with other checkpoint-based interventions in Australia and world wide, the "secret" of this apparent success would seem to lie in the quantum increase in enforcement and publicity resources initially invested in the program. RBT was also favored with the luck of being introduced coincident with an economic recession, which magnified the subsequent reduction in deaths and injuries and enhanced perception of the program as effective.

The New South Wales experience has not been replicated in the other Australian states which, with the exception of Tasmania, until recently were far more sparing of resources invested in checkpoint-based programs. The Tasmanian experience has not been evaluated, although the testing ratio there is even higher than that in New South Wales. It certainly would seem to

merit study, along with the results of an intense RBT program adopted in the state of Victoria in late 1990 (Gerondeau, 1991).

In considering the import of the New South Wales experience for the U.S., it is important to note the difference between RBT and sobriety checkpoints as we have experienced them. Our programs have a far greater resemblance to those in the other Australian states, the evaluations of which have shown either a lack of effect or one that fades very quickly from a promising inception. One source of limitations on what we do may be principle, respecting the fact that "the costs in a democratic society of a system of enforcement and a style of publicity which rely increasingly on the creation of feelings of terror in the driving public should not be underemphasized" (Homel, 1988: 271). Although one may find the principle outweighed by activity that saves lives, one cannot dismiss it out of hand. But another source of limitations may be reluctance to make the massive commitment of resources to prevention that seems to have attained important and long-lasting results in New South Wales.

Conclusion

Both U.S. and foreign experiences provide support for the proposition that sobriety checkpoints are capable of reducing the extent of drunk driving and of deaths and injuries on the highways. It is no longer necessary to ask whether sobriety checkpoints have deterrent potential. However, one may and should still inquire into the conditions under which they work well, resulting in a steep and long-term decline in fatalities, or poorly, possibly representing a waste of resources compared with other possible uses for them. The following comments represent speculation based on experience to date.

A first consideration would seem to be the nature of the checkpoints. It must be obvious to drivers passing through them, and ideally to drivers passing by, that detecting drunk driving is the main point. In Honolulu's otherwise impressive checkpoint program, drivers are informed by signs merely that a "checkpoint" is ahead. When they are stopped, they are told that they are in a "holiday checkpoint" and are asked to produce their licenses, vehicle registration, and insurance papers. They may be cited for failing to present these or for other violations, such as defective equipment, that become apparent in the check. Although the police are alert for cues to alcohol and drug impairment and will arrest suspected drunk drivers, the passing public is not routinely informed that this inquiry is a principal or even a subordinate goal of the enterprise. In New York City, sobriety checkpoints may have no signs at all communicating their purpose to passing motorists, and many drivers are waved through without even brief interviews. It is hard to believe that programs like these can do much to increase the perceived risk that a drunk

driver will be apprehended. In contrast, in checkpoints operated by the Delaware State Police and in certain communities in California and Pennsylvania, drivers are informed that a "sobriety checkpoint" is ahead, and this is repeated by signs when they are stopped as well as by the officers addressing them. No documents are requested unless a brief interview suggests the likelihood of drunk driving. Not only is the purpose of the checkpoint made evident to all who pass by and can read, but the police themselves are clear concerning their goals. The procedure is also minimally intrusive and thus of comparatively reduced concern from a civil liberties perspective.

A second consideration would seem to be the frequency of the checkpoints. Although deterrent results have been found from programs involving as few as two per site, they were fleeting. Long-term results might most reasonably be expected from long-term commitment to a checkpoint program, or at least to an enforcement program in which checkpoints play a part. The checkpoints must be frequent, perhaps exceeding a threshold beneath which they can be regarded as unthreatening. It is possible that the Indianapolis checkpoint program may have been compromised because of insufficient frequency. The most credible deterrence claims in the world literature concern programs where the numbers of drivers tested constitute a substantial fraction of the numbers of drivers licensed in the jurisdiction.

Third, effective checkpoints seem to require publicity, regardless of the level of enforcement. Although novel programs employing checkpoints are generally newsworthy, the New Mexico experience suggests that this is not inevitable and also that unpublicized enforcement gets few results. The public was able to disregard totally a major checkpoint program in British Columbia when the media were silenced by a strike. In contrast, the effective New South Wales program was launched with enormous amounts of publicity, much of it paid. The new California programs have obtained continued attention over many months by providing the media with releases stating the time and place of checkpoints two hours before they take place. If a few impaired but attentive drivers use this information to avoid detection, this may be a small price for reminding the public of the existence of the checkpoint effort.

A fourth issue is the relationship between sobriety checkpoints and other forms of drunk-driving law enforcement. Not only must we learn more about the comparative effectiveness of checkpoints and other forms of patrol -- although both can deter drunk driving, their relative efficiency and effectiveness have not been systematically studied -- but we must learn what part checkpoints can play in an enforcement strategy that includes other elements as well. It is possible that an ideal strategy for deterring drunk drivers will contain elements of traditional patrol in high-incidence areas and times, novel patrol practices

emphasizing testing of participants in accidents throughout the jurisdiction, and sobriety checkpoints at high-visibility times and places.

References

Arthurson, R. 1985. Evaluation of Random Testing. Sydney: Traffic Authority of New South Wales.

Cameron, M., Strang, P., and Vulcan, A. 1980. Evaluation of random breath testing in Victoria, Australia. Paper presented at the Eighth International Conference on Alcohol, Drugs and Traffic Safety, Stockholm.

Carseldine, D. 1985. Surveys of Knowledge, Attitudes, Beliefs and Reported Behaviours of Drivers -- on the Topic of Drink-Driving and Random Breath Testing. Sydney: Traffic Authority of New South Wales.

Dickman, F., 1987. Sobriety checkpoints for DWI enforcement: A review of current research. Unpublished manuscript, Traffic Safety Programs, National Highway Traffic Safety Administration.

Epperlein, T. 1985. The use of sobriety checkpoints as a deterrent: An impact assessment. Report from Arizona Department of Public Safety, Phoenix.

Evans, W., Neville, D., and Graham, J. 1991. General deterrence of drunk driving: Evaluation of recent American policies. Risk Analysis 11: 279-289.

Gerondeau, Christian. 1991. La Securite Routiere en Australie. Paris: ADTS.

Hingson, R., and Howland, J. 1989. Public perception about drunk-driving roadblocks: Surveys in Massachusetts and New Hampshire. Report to Massachusetts Governor's Highway Safety Bureau.

Homel, R. 1988. Policing and Punishing the Drinking Driver: A study of General Deterrence. New York: Springer-Verlag.

Homel, R., 1990. Random breath testing and random stopping programs in Australia. In Wilson, R.J., and Mann, R. (eds.), Drinking and Driving: Advances in Research and Prevention. New York: Guilford Press, pp. 159-204.

Joksch, H. 1988. The Impact of Severe Penalties on Drinking and Driving. Washington, DC: AAA Foundation for Traffic Safety.

Klein, T, 1989. Changes in Alcohol-Involved Fatal Crashes Associated with Tougher State Alcohol Legislation. Technical

report. Washington, DC: National Highway Traffic Safety Administration.

Lacey, J., Marchetti, L., Stewart, J., Popkin, C., Murphy, P., and Jones, R. 1990. Combining Enforcement and Public Information to Deter DWI" The Experience of Three Communities. Technical report. Washington, DC: National Highway Traffic Safety Administration.

Lacey, J., Marchetti, L., Stewart, J., Popkin, C., Murphy, P., Lucke, R., Jones, R., and Ruschmann, P. 1988. Enforcement and Public Information Strategies for DWI General Deterrence: The Indianapolis, Indiana Experience. Technical report. Washington, DC: National Highway Traffic Safety Administration.

Lacey, J., Stewart, J., Marchetti, L., Popkin, C., Murphy, P., Lucke, R., and Jones, R. 1986. Enforcement and Public Information Strategies for DWI General Deterrence: Arrest Drunk Driving: The Clearwater and Largo, Florida, Experiences. Technical report. Washington, DC: National Highway Traffic Safety Administration.

Levy, D., Shea, D, and Asch, P., 1988. Traffic safety effects of sobriety checkpoints and other local DWI programs in New Jersey. American Journal of Public Health 79: 291-293.

Levy, D., Asch, P., and Shea, D., 1990. An assessment of county programs to reduce driving while intoxicated. Health Education Research 5: 247-255.

Mercer, G. W., 1984. Alcohol-Related Casualty Traffic Accidents in British Columbia. Vancouver, Canada: Ministry of Attorney General.

Mercer, G. W., 1985. The relationships among driving while impaired charges, police drinking-driving roadcheck activity, media coverage and alcohol-related casualty traffic accidents. Accident Analysis and Prevention 17: 467-474.

Michigan Department of State Police v. Sitz. 1990. 110 S. Ct. 2481.

NHTSA (National Highway Traffic Safety Administration) 1990. The Use of Sobriety Checkpoints for Impaired Driving Enforcement. Technical report. Washington, DC: NHTSA Office of Enforcement and Emergency Services.

NRMA (National Roads and Motorists Association). 1985. NRMA Submission to the NSW Standing Committee on Road Safety Part 5: Random Breath Testing. Sydney: NRMA.

NRMA (National Roads and Motorists Association). 1991. The Role of Legal and Illegal Drugs in Road Crashes: NRMA Submission to the NSW Standing Committee on Road Safety, Part 11. Sydney: NRMA.

RTA (Roads and Traffic Authority). 1991. Road Traffic Accidents in NSW -- 1990. Sydney: RTA.

Ross, H. L. 1982. Detering the Drinking Driver: Legal Policy and Social Control. Lexington, MA: D. C. Heath.

Ross, H. L., McCleary, R., and Epperlein, T. 1982. Deterrence of drinking and driving in France: An evaluation of the law of July 12, 1978. Law and Society Review 16: 345-374.

Ross, H. L., and Voas, R., 1989. The New Philadelphia Story: The Effects of Severe Penalties for Drunk Driving. Washington, DC: AAA Foundation for Traffic Safety.

Voas, R., Rhodenizer, E., and Lynn, C. 1985. Evaluation of Charlottesville Checkpoint Operations: Final Report. Technical report. Washington, DC: National Highway Traffic Safety Administration.

Wells, J., Preusser, D., and Williams, A. 1991. Enforcing alcohol-impaired driving and seat belt use laws, Binghamton, New York. Report from Insurance Institute for Highway Safety, Arlington, VA.

Williams, A., and Lund, A. 1984. Deterrent effects of roadblocks on drinking and driving. Traffic Safety Evaluation Research Review 3: 7-18.

Zador, P., Lund, A., Fields, M., and Weinberg K. 1988. Fatal crash involvement and laws against alcohol-impaired driving. Journal of Public Health Policy 10: 467-485.

Zimring, F., and Hawkins, G. 1973. Deterrence: The Legal Threat in Crime Control. Chicago: University of Chicago Press.

Table 1. Summary of the Literature on Sobriety Checkpoints in the United States

<u>Site and Author</u>	<u>Nature of Intervention</u>	<u>Effect on Perception</u>	<u>Effect on Impaired Driving</u>	<u>Comments</u>
Maryland, Delaware (Williams & Lund 1984)	49 ckpts./yr. in 2 DE counties 30-50 ckpts./year for 2 yrs. in Montgomery Co., MD Fri. and Sat. nights	Greater awareness of D/D activities, higher perceived risk of arrest, than in control counties	No effect according to self reports	Fairly good quasi-experimental design
Arizona (Epperlein 1985)	2 ckpts. in each of 3 sites on a December weekend	No data	Drop of 29% in ratio of fatal and injury crashes to total crashes, by ITS method. 98% gone in 2 mos. Possible instrumentation problem	Very modest intervention well analyzed
19 Charlottesville (Voas et al 1985)	94 ckpts., 24,000 tests Fri. and Sat. nights for one year.	Interviews with alcohol servers; tel. survey with base and post-ckpt. waves, in Charlottesville and control. Changes in expos. and knowledge, not in perceived risk	Before-after comparison finds decline in 6 measures, only 3 signif. Control with balance of State	Fair method due to short data series
Clearwater Largo (Lacey et al 1986)	12 ckpts. in 15 mos. "Media events" Complex intervention includes wolfpacks	Interviews in site and control areas find greater knowledge of changed enforcement, higher perceived risk of punishment	ITS finds reduction in alc. related crashes and night-time crashes	Complex intervention clouds causal relationships
Indianapolis (Lacey et al 1988)	6 ckpts. with extensive news coverage	Ckpts. perceived, but no change in perceived risk	ITS finds no significant change	Few ckpts., complex intervention
Massachusetts (Hingson & Howland 1989)	MA police "use" ckpts., whereas NH police cannot	Ckpts. are most frequently mentioned police activity, are believed more effective than patrols	No data	Vague intervention no direct measure of effect

Table 1. Summary of the Literature on Sobriety Checkpoints in the United States (cont.)

<u>Site and Author</u>	<u>Nature of Intervention</u>	<u>Effect on Perception</u>	<u>Effect on Impaired Driving</u>	<u>Comments</u>
Bergen Co. (Levy et al 1990 and other papers)	35,000 tests/yr. for 2-1/2 yr., weekend nights. Separate program generated publicity	No data	10-15% decline in a single vehicle night accidents, using ITS. Effect peaks in 2nd year and declines. Not confirmed in nighttime fatality data.	Contestable assumptions, inconsistency of measured effects
Binghamton (IIHS, 1990)	72 ckpts. in 2 year, 9,400 tests	Awareness, perceived risk peak following inception, compared with control city	Decline in % alcohol +, but not in BAC over .10%. Sig. declines in injury and late night crashes using ITS	Strong quasi- experimental design
New Mexico (Sewell, forth- coming)	4 sites: a. 101 ckpts., 30,000 tests b. 18 ckpts., 1,300 tests c. 28 ckpts., 2,300 tests d. 34 ckpts., 12,000 tests 2+ yr. for (a), 1 yr. others	No data; publicity appears weak	Fatal + inj. crashes down 11% in 1st year, up 11% in 2nd. In site (a) only, down 14%, up 21. B-A comparison, no stat. tests	Analysis incomplete to date

Monthly Count of Fatal Crashes In New South Wales January 1981 to March 1985

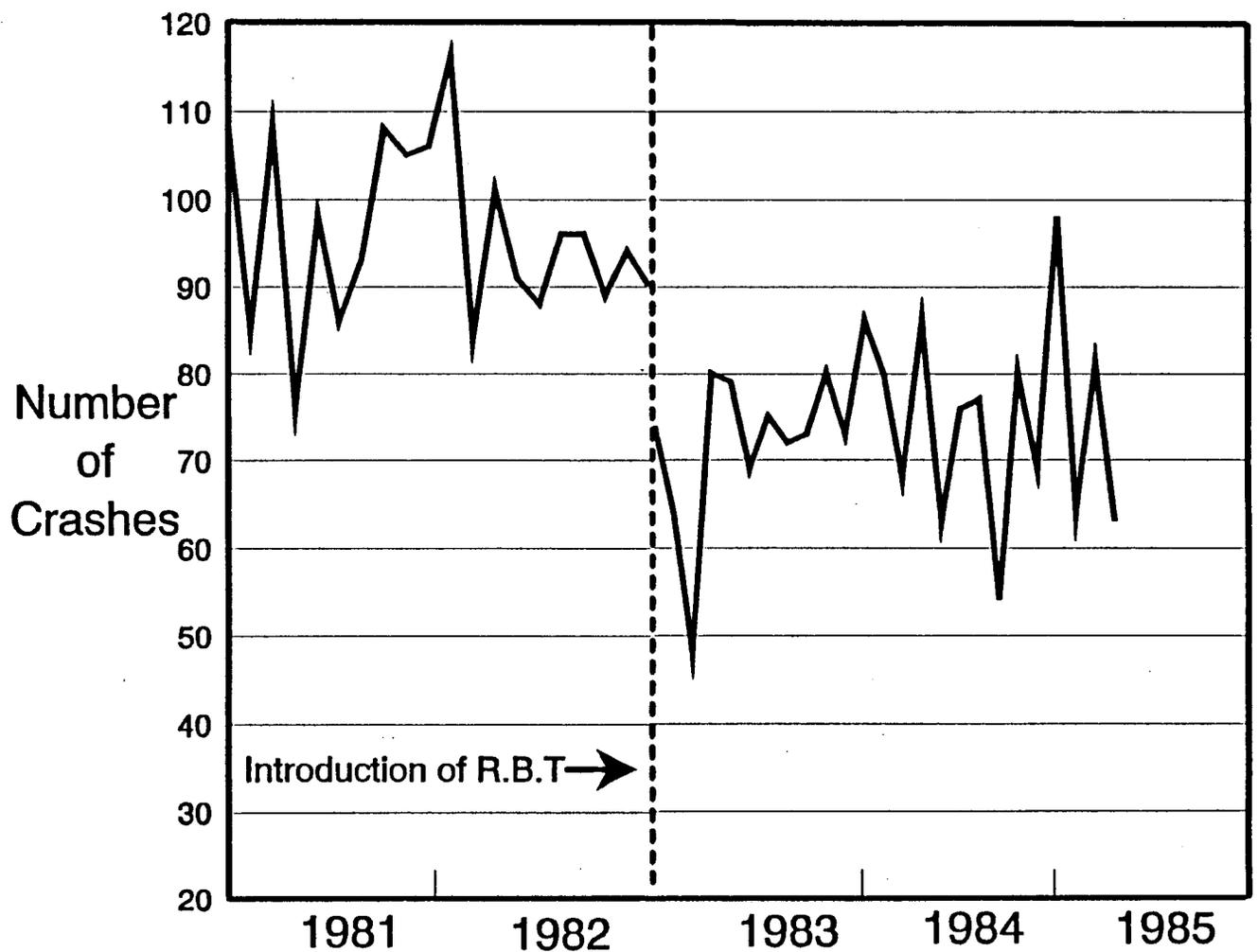


Figure 1

Source: Arthurson, 1985